ELSEVIER

Stochastic Processes and their Applications 93 (2001) 317–337

stochastic processes and their applications

www.elsevier.com/locate/spa

A strong invariance principle for the logarithmic average of sample maxima ☆

Ingo Fahrner

Abteilung Zahlentheorie und Wahrscheinlichkeitstheorie, Universität Ulm, 89069 Ulm, Germany

Received 18 April 2000; received in revised form 20 November 2000; accepted 27 November 2000

Abstract

Given an extremal-Λ process $\{Y_\Lambda(t),\ t>0\}$, the transformed process $\{U(s)=Y_\Lambda(e^s)-s,\ -\infty<s<\infty\}$ is a stationary strong Markov process. We prove an almost sure invariance principle for the process $\{\int_0^t f(U_s)\,ds,\ t\geqslant 0\}$. By an approximation this yields an almost sure invariance principle for the logarithmic average of normed sample maxima, which have been investigated recently in various papers. With this invariance principle, we can also get various results on the behavior of sums of minima of a sequence of random variables.

MSC: primary 60F17; 60G70; secondary 60F05; 60F15

Keywords: Wiener process; Extremal process; Strong approximation; Invariance principle; Almost sure behavior of extremes

1. Introduction

Let $X_1, X_2, \ldots$ be independent, identically distributed (i.i.d.) random variables with common distribution function F and set $M_n := \max_{1\leqslant j\leqslant n} X_j$ for $n\in\mathbb{N}$. If there exist sequences of real numbers $\{a_n\}$ and $\{b_n\}$ with $a_n>0$ for all $n\in\mathbb{N}$ such that

$$\lim_{n\to\infty} P((M_n-b_n)/a_n\leqslant x)=G(x) \quad \text{for all continuity points of } G \tag{1}$$

for some non-degenerate distribution function G then G is of extreme value type and F is said to belong to the domain of attraction of G ($F\in\mathscr{D}(G)$). More precisely, there exist $a>0$ and $b\in\mathbb{R}$ such that $G(ax+b)$ equals one of the following three distribution functions:

$$\Phi_\alpha(x) = \exp(-x^{-\alpha})I_{(0,\infty)}(x) \quad \text{for some } \alpha>0,$$

$$\Psi_\alpha(x) = \exp(-(-x)^\alpha)I_{(-\infty,0]}(x)+I_{(0,\infty)}(x) \quad \text{for some } \alpha>0$$

☆ This paper is a part of the author's dissertation thesis written at the University of Ulm under the direction of Professor Ulrich Stadtmüller. The author wants to express his sincere gratitude to Professor Stadtmüller for his guidance.

E-mail address: fahrner@mathematik.uni-ulm.de (I. Fahrner).

0304-4149/01/$ - see front matter
PII: S0304-4149(00)00103-4

or

$$\Lambda(x) = \exp(-e^{-x}).$$

There is a canonical choice for the norming sequences $\{a_n\}$ and $\{b_n\}$ and in this paper we will restrict ourselves to these canonical choices (see below). Further information on extreme value distributions, their domains of attraction and norming constants can be found in the books of de Haan (1970), Leadbetter et al. (1983) or Resnick (1987).

Concerning the almost sure behavior of $(M_n - b_n)/a_n$ Fahrner and Stadtmüller (1998) and independently Cheng et al. (1998) prove that if (1) holds then

$$\frac{1}{\log n}\sum_{k=1}^{n}\frac{1}{k}f((M_k - b_k)/a_k) \to \int_{\mathbb{R}} f(x)\,\mathrm{d}G(x) \quad \text{a.s. as } n\to\infty \tag{2}$$

for every almost everywhere continuous and bounded function f. In Fahrner (2000a, b), it is shown that the assumption of boundedness of f can be weakened considerably and can be replaced by a condition which demands somewhat more than the existence of the integral on the right-hand side of (2).

In this paper, we want to study the rate of convergence and the asymptotic distribution of logarithmic averages as in (2) for a large class of functions f. The results parallels those for partial sums instead of maxima: Multiplying the logarithmic average in (2) by $\sqrt{\log n}$ we get asymptotic normality. In fact, we can show a strong invariance principle with a rate which is good enough to imply the law of the iterated logarithm and a functional central limit theorem.

There is an analogous theory for logarithmic averages of partial sums. See Berkes (1998) for a detailed survey and for references.

2. Main results

We first define the class of functions f we will work with. Examples and properties of these functions will be given at the end of this section.

Let $\mathscr{M}_\varrho$, $\varrho\in[0,\frac{1}{2})$ denote the set of finite linear combinations of functions $f:\mathbb{R}\to\mathbb{R}$ which satisfy

$$\sup_{x\in\mathbb{R}}|f(x)| < \infty \quad \text{and} \quad f \text{ is non-decreasing} \tag{3}$$

or

$$\sup_{x\in\mathbb{R}}\sup_{s>0}\frac{|f(x+s)-f(x)|}{h_\gamma(x)(e^{\gamma s}-1)} < \infty \quad \text{for some } 0<\gamma<\tfrac{1}{2}-\varrho, \tag{4}$$

where $h_\gamma(x) = e^{\gamma x} + \exp(e^{-\gamma x})e^{-\gamma x}$.

We denote by $\bar{F}(x) := 1 - F(x)$ the tail of F and by $x_F := \sup\{x\in\mathbb{R}: F(x) < 1\}$ the right end of the distribution F. If $F\in\mathscr{D}(G)$ has a second derivative we let $a(x) := \bar{F}(x)/F'(x)$ and define $\gamma_k := \inf\{x\in\mathbb{R}: 1/\bar{F}(x)\geqslant k\}$ and the canonical choices for a_k and b_k by

$$a_k = a(\gamma_k), \quad b_k = \gamma_k \quad \text{when } G = \Lambda,$$

$$a_k = \gamma_k, \quad b_k = 0 \quad \text{when } G = \Phi_\alpha,$$

$$a_k = x_F - \gamma_k, \quad b_k = x_F \quad \text{when } G = \Psi_\alpha.$$

Our first result is

Theorem 1. *Let $\{X_n,\ n \in \mathbb{N}\}$ be a sequence of i.i.d. random variables with common distribution function F. Suppose F has a negative second derivative in some neighborhood of x_F and for $a(x) := \bar{F}(x)/F'(x)$ we have*

$$\lim_{x \to x_F} a'(x)(\log\log(1/\bar{F}(x)))^2 (\log(1/\bar{F}(x)))^{1/2+2\eta} = 0 \tag{5}$$

for some $\eta > 0$. Let a_k and b_k, $k \in \mathbb{N}$, be the canonical choices of norming constants.

Then we can redefine the sequence $\{X_n,\ n \in \mathbb{N}\}$ together with an extremal-Λ process $\{Y_\Lambda(u),\ u > 0\}$ on a richer probability space without changing the distribution of the sequence such that

$$\left| \sum_{k=1}^{n} \frac{1}{k} f\left(\frac{M_k - b_k}{a_k}\right) - \int_1^n \frac{1}{u} f(Y_\Lambda(u) - \log u)\,\mathrm{d}u \right| = \mathrm{o}((\log n)^{1/2-\eta}) \quad a.s. \tag{6}$$

as $n \to \infty$ for every $f \in \mathscr{M}_0$.

Note that by (5) the von Mises condition $\lim_{x \to x_F} a'(x) = 0$ holds and therefore $F \in \mathscr{D}(\Lambda)$. Condition (5) is an adaptation of the de Haan and Hordijk (1972) condition and is necessary to couple the behaviors of $(M_k - b_k)/a_k$ and Y_Λ. In general, it is not possible to construct an extremal-Λ process which is close to $(M_k - b_k)/a_k$ without condition (5) (see Fahrner and Stadtmüller, 2000). In applications, (5) is rather mild: If F is the normal distribution or a Weibull distribution $\bar{F}(x) = \exp(-cx^{-\tau})$ for $c, \tau, x > 0$ (including the exponential distribution) we may take any $0 < \eta < \frac{1}{4}$.

Our second result is

Theorem 2. *Let $\{Y_\Lambda(t),\ t > 0\}$ be an extremal-Λ process and define $U(s) = U_s := Y_\Lambda(\mathrm{e}^s) - s$ for $s \in \mathbb{R}$. For every $f \in \mathscr{M}_\varrho$, $0 < \varrho < \frac{1}{4}$, it is possible to redefine the process U on a richer probability space without changing its distribution together with a Wiener process $\{W(t),\ t \geqslant 0\}$ such that*

$$\left| \int_0^t (f(U_s) - m)\,\mathrm{d}s - \sigma W(t) \right| = \mathrm{o}(t^{1/2-\varrho}) \quad a.s. \text{ as } t \to \infty, \tag{7}$$

where $m = \int_{-\infty}^{\infty} f(x)\,\mathrm{d}\Lambda(x)$ and

$$\sigma^2 = 2\int_0^\infty \left(\int_{-\infty}^{\infty} f(y) f(y-r) \Lambda(y-r) \mathrm{e}^{-y}\,\mathrm{d}y \right.$$
$$\left. + \int_{-\infty}^{\infty} \int_{y-r}^{\infty} f(x) f(y) (1 - \mathrm{e}^{-r}) \Lambda^{1-\mathrm{e}^{-r}}(x) \mathrm{e}^{-x} \Lambda(y) \mathrm{e}^{-y}\,\mathrm{d}x\,\mathrm{d}y - m^2 \right) \mathrm{d}r. \tag{8}$$

Theorems 1 and 2 can be combined since by the substitution $s = \log u$ we have

$$\int_1^n \frac{1}{u} f(Y_\Lambda(u) - \log u)\,\mathrm{d}u = \int_0^{\log n} f(U_s)\,\mathrm{d}s.$$

Thus, we obtain the following corollaries.

Corollary 3. *Let $\{X_n,\ n \in \mathbb{N}\}$ be a sequence of i.i.d. random variables with common distribution function $F \in \mathscr{D}(G)$ and g be a function. Let*

$$f(x) = g(x), \quad F^{\#}(x) = F(x) \quad \text{when } G = \Lambda,$$

$$f(x) = g(\mathrm{e}^{x/\alpha}), \quad F^{\#}(x) = F(\mathrm{e}^x) \quad \text{when } G = \Phi_\alpha,$$

$$f(x) = g(-\mathrm{e}^{-x/\alpha}), \quad F^{\#}(x) = F(x_F - \mathrm{e}^{-x}) \quad \text{when } G = \Psi_\alpha,$$

$\{a_k\}$ and $\{b_k\}$ be the canonical choices of norming sequences and define m and σ^2 as in Theorem 2. Suppose $F^{\#}$ has a negative second derivative in some neighborhood of $x_{F^{\#}}$, satisfies (5) for some $0 < \eta < \frac{1}{4}$ and if $G = \Phi_\alpha$ or $G = \Psi_\alpha$ we have in addition

$$\lim_{x\to\infty} (\alpha \cdot a(x) - 1)x^{1/2+2\eta} \log x = 0. \tag{9}$$

Then if $f \in \mathscr{M}_\eta$ we can redefine the sequence $\{X_n,\ n \in \mathbb{N}\}$ together with a Wiener process $\{W(t),\ t \geqslant 0\}$ on a richer probability space such that

$$\sum_{k=1}^n \frac{1}{k} g\left(\frac{M_k - b_k}{a_k}\right) - m \log n = \sigma W(\log n) + \mathrm{o}((\log n)^{1/2-\eta}) \quad a.s. \tag{10}$$

as $n \to \infty$.

Remark 4. Conditions similar to (5) can be given for $F \in \mathscr{D}(\Phi_\alpha)$ or $F \in \mathscr{D}(\Psi_\alpha)$ directly, see Theorems 4 and 5 of Fahrner and Stadtmüller (2000).

For the logarithmic average of maxima we end with the following results.

Corollary 5 (Functional central limit theorem). *Under the assumptions of Corollary* 3, *if $f \in \mathscr{M}_0$ then*

$$\frac{1}{\sqrt{\log n}} \left(\sum_{k=1}^{[n^t]} \frac{1}{k} g\left(\frac{M_k - b_k}{a_k}\right) - mt \log n \right) \xrightarrow{\mathrm{d}} \sigma W(t) \quad \text{in } D[0,1],$$

where $\{W(t),\ 0 \leqslant t \leqslant 1\}$ is a Wiener process.

Corollary 6 (Laws of the iterated logarithm). *Under the assumptions of Corollary* 3, *if $f \in \mathscr{M}_0$ then*

$$\limsup_{n\to\infty} \frac{1}{\sqrt{2 \log n \log\log\log n}} \max_{1 \leqslant j \leqslant n} \left| \sum_{k=1}^j \frac{1}{k} g\left(\frac{M_k - b_k}{a_k}\right) - m \log j \right| = \sigma \quad a.s.,$$

$$\liminf_{n\to\infty} \sqrt{\frac{\log\log\log n}{\log n}} \max_{1 \leqslant j \leqslant n} \left| \sum_{k=1}^j \frac{1}{k} g\left(\frac{M_k - b_k}{a_k}\right) - m \log j \right| = \frac{\pi}{\sqrt{8}} \sigma \quad a.s.$$

Examples for functions belonging to $\mathscr{M}_\varrho$, $\varrho \in [0, \frac{1}{2})$, are indicators of finite unions of intervals or functions of bounded variation with compact support. Functions satisfying (4) are locally Lipshitzian with a Lipshitz constant which may grow with a certain rate. In fact every Lipshitz function is in $\mathscr{M}_\varrho$.

For differentiable functions f, by the mean value theorem, a sufficient condition for $f \in \mathscr{M}_\varrho$ is that for some $C > 0$, $0 < \gamma < \frac{1}{2} - \varrho$ and any $s > 0$

$$\sup_{x \leqslant \xi \leqslant x+s} |f'(\xi)| \leqslant C e^{\gamma s} h_\gamma(x)$$

holds. Thus, the following functions belong to $\mathscr{M}_\varrho$: $|x|^r$, $r \geqslant 1$, $e^{\gamma x}$, $\exp(e^{-\gamma x})$, $0 < \gamma < \frac{1}{2} - \varrho$, $\sin x$, $\cos x$, etc.

The functions $|x|^r$, $0 < r < 1$ are not Lipshitz continuous, but since these functions can be written as sums of a Lipshitz continuous function and a function of bounded variation with compact support, combining (3) and (4) we see that they nevertheless belong to $\mathscr{M}_\varrho$.

It can be seen easily that if $f \in \mathscr{M}_\varrho$ then there is some $C > 0$ and some $0 < \gamma' < \frac{1}{2} - \varrho$ with

$$|f(x)| \leqslant C h_{\gamma'}(x) \tag{11}$$

and thus $\int_{\mathbb{R}} f^{2/(1-2\varrho)}(x)\, d\Lambda(x) < \infty$: If f satisfies (4) with some γ then setting $x = 0$ in (4) yields $|f(s)| \leqslant C e^{\gamma s}$ for all $s \geqslant 0$ and setting $s = -x$ for $x < 0$ in (4) yields $|f(x)| \leqslant C h_\gamma(x) e^{-\gamma x} \leqslant C h_{\gamma'}(x)$ with some γ' which is slightly larger than γ.

We finish this section with a discussion on related work by Berkes and Horváth (2001) which we received during the preparation of this manuscript.

Let $Z(n) = \sum_{k=1}^{n} 1/k f((M_k - b_k)/a_k)$, then Berkes and Horváth (2001) prove that if f is of bounded variation with compact support there is a Wiener process $\{W(t),\ t \geqslant 0\}$ such that

$$Z(n) - \mathbb{E}Z(n) = \sigma W(\tau_n) + o((\log n)^{1/2-\varepsilon}) \quad \text{a.s. as } n \to \infty, \tag{12}$$

where $\tau_n \sim \log n$, $0 < \varepsilon < \frac{1}{24}$ and σ is given by (8). Our result needs the second-order condition (5) and reads

$$Z(n) - m \log n = \sigma W(\log n) + o((\log n)^{1/2-\eta}) \quad \text{a.s. as } n \to \infty, \tag{13}$$

where m is given in Theorem 2 and $0 < \eta < \frac{1}{4}$ depends on the underlying distribution and the choice of f. We allow more general functions f, which might be unbounded. If f is of bounded variation with compact support then clearly $\mathbb{E}Z(n) - m \log n \to 0$ as $n \to \infty$, but in general the speed of convergence in this relation is not clear. Thus, neither (12) follows from (13), nor (13) from (12).

The idea of proof in Berkes and Horváth (2001) is the following: First, $Z(n)$ is divided in blocks $\beta_1, \eta_1, \beta_2, \eta_2, \ldots$ and it is shown that the blocks $\eta_1, \eta_2, \ldots$ can be neglected. The blocks β_i can be approximated by other blocks β_i^* for which an invariance principle of Strassen applies. The advantage of this method is that it applies without any second-order condition on the underlying distribution. However, the boundedness of f is essential for the above mentioned approximations.

Turning to our method, note that with the embeddings of Corollary 3 it suffices to consider the case $F \in \mathscr{D}(\Lambda)$. Here, the problem of a strong approximation for $Z(n)$ can be reduced to the problem of giving a strong approximation for $\int_0^t f(U_s)\,\mathrm{d}s$, where U_s is an explicitly given process which is the same for all $F\in\mathscr{D}(\Lambda)$. This time continuous problem can be handled very well for a wide class of functions f, even when they are unbounded. However, for the reduction (Theorem 1) we need condition (5).

3. Connections to the study of sums of minima

Let $X_1,X_2,\dots$ be i.i.d. random variables and set $m_n:=\min\{X_1,\dots,X_n\}$ and $S_n:=\sum_{k=1}^n m_k,\ n\in\mathbb{N}$. Many authors investigated the asymptotic behavior of S_n: Grenander (1965) proves a weak law, strong laws are proved by Deheuvels (1974), Gosh et al. (1975) and Gouet (1989), Höglund (1972) and Deheuvels (1974) show asymptotic normality and invariance principles are given by Hebda-Grabowska and Szynal (1979), Gouet (1989) and Gladysheva and Sakhanenko (1985).

In all proofs, one first assumes that X_1 has a uniform distribution on $(0,1)$ and then a quantile transformation shows the general result.

Our results give a new proof for the case when X_1 has a uniform distribution on $(0,1)$. The uniform distribution is in the domain of attraction of Ψ_1 and

$$k(M_k-1)\xrightarrow{\mathrm{d}}\Psi_1 \quad \text{as } k\to\infty.$$

We have $\{X_k,k\in\mathbb{N}\}\overset{\mathrm{d}}{=}\{1-X_k,\ k\in\mathbb{N}\}$ in $\mathbb{R}^\infty$ and thus $\{m_k,\ k\in\mathbb{N}\}\overset{\mathrm{d}}{=}\{-(M_k-1),\ k\in\mathbb{N}\}$ in $\mathbb{R}^\infty$, i.e.

$$\{S_n,\ n\in\mathbb{N}\}\overset{\mathrm{d}}{=}\left\{\sum_{k=1}^n\frac{1}{k}g(k(M_k-1)),\ n\in\mathbb{N}\right\}\quad\text{in } \mathbb{R}^\infty,$$

where $g(x)=-x$. Applying Corollary 3 we get $f(x)=\mathrm{e}^{-x}$, $\overline{F^{\#}}(x)=\mathrm{e}^{-x}$ and $a(x)\equiv 1$, therefore $m=1$ and $\sigma^2=2$ and for every $\varepsilon>0$ there is a Wiener process $\{W_t,\ t\geqslant 0\}$ such that

$$S_n-\log n=W(2\log n)+\mathrm{o}((\log n)^{1/4+\varepsilon})\quad\text{a.s. as } n\to\infty.$$

The rate here is better than that of Hebda-Grabowska and Szynal (1979), who get $\mathrm{o}((\log n)^{3/8+\varepsilon})$, but not optimal. Gladysheva and Sakhanenko (1985) give a construction with the error $\mathcal{O}(\log\log n)$.

4. Proofs

Proof of Theorem 1. Let $I(A)$ denote the indicator function of the set A and set $\log^* u=\log(\max\{e,u\})$, $\log_3^* u=\log^*\log^*\log^* u$. By Theorem 3 of Fahrner and Stadtmüller (2000), we can redefine the sequence $\{X_n,\ n\in\mathbb{N}\}$ together with an extremal-Λ process $Y_\Lambda\in\mathscr{D}(0,\infty)$ on a richer probability space (Ω,Σ,P) without changing the distribution of the sequence such that given $\varepsilon>0$ there is an event $\Omega_\varepsilon\in\Sigma$ with $P(\Omega_\varepsilon)\geqslant 1-\varepsilon$ and a non-negative function $r_\varepsilon(u)$ such that for some $K_\varepsilon\geqslant 1$

$$Y_\Lambda(u-v(u))-\log u-r_\varepsilon(u)\leqslant\frac{M_{[u]}-b_{[u]}}{a_{[u]}}\leqslant Y_\Lambda(u+v(u))-\log u+r_\varepsilon(u) \qquad (14)$$

for all $u \geqslant K_\varepsilon^2$ and all $\omega \in \Omega_\varepsilon$. Here $v(u) = K_\varepsilon \sqrt{\log^* u \log_3^* u}$ and (5) implies $r_\varepsilon(u)(\log u)^{1/2+2\eta} \to 0$ as $u \to \infty$. Recall that the constructed process $Y_A(u)$ does not depend on the choice of ε. We define

$$T(u) := \frac{M_{[u]} - b_{[u]}}{a_{[u]}} \quad \text{and} \quad Z(u) := Y_A(u) - \log u \quad \text{for } u \geqslant 1.$$

We will show that for all $\varepsilon > 0$

$$\frac{1}{(\log n)^{1/2-\eta}} \left| \sum_{k=1}^{n} \frac{1}{k} f(T(k)) - \int_1^n \frac{1}{u} f(Z(u))\,\mathrm{d}u \right| \to 0 \quad \text{as } n \to \infty$$

for almost all $\omega \in \Omega_\varepsilon$. Since $\varepsilon > 0$ is arbitrary, (6) will follow.

Fix $\varepsilon > 0$. We have

$$\begin{aligned}
\sum_{k=1}^{n-1} \frac{1}{k} f(T(k)) &= \sum_{k=1}^{n-1} \int_k^{k+1} \frac{1}{k} f(T(k))\,\mathrm{d}u \\
&= \sum_{k=1}^{n-1} \int_k^{k+1} \left(\frac{1}{k} - \frac{1}{u}\right) f(T(k))\,\mathrm{d}u + \int_1^n \frac{1}{u} f(T(u))\,\mathrm{d}u \qquad (15)
\end{aligned}$$

by addition and subtraction of

$$\int_1^n \frac{1}{u} f(T(u))\,\mathrm{d}u = \sum_{k=1}^{n-1} \int_k^{k+1} \frac{1}{u} f(T(k))\,\mathrm{d}u.$$

Note that by assumption there is a constant $C_1 > 0$ such that $|f(x)| \leqslant C_1 \exp(\exp(|x|/2))$ and condition (5) implies that for almost every ω there is a $k_0(\omega) \in \mathbb{N}$ such that for $k \geqslant k_0$ we have $|T(k)| \leqslant \frac{3}{2} \log\log k$ (cf. Theorem 2 of de Haan and Hordijk, 1972). Thus, using $x^{3/4} \leqslant (3x+1)/4$ for $x \geqslant 0$, we see

$$|f(T(k))| \leqslant C_1 \exp((\log k)^{3/4}) \leqslant C_1 \mathrm{e}^{1/4} k^{3/4} \quad \text{for } k \geqslant k_0(\omega)$$

and the first term in (15) can for almost all ω be bounded by

$$\sum_{k=1}^{n-1} \left(\frac{1}{k} - \frac{1}{k+1}\right) |f(T(k))| \leqslant C_2(\omega) \sum_{k=1}^{\infty} k^{-5/4} < \infty.$$

Therefore, almost surely

$$\begin{aligned}
\left| \sum_{k=1}^{n} \frac{1}{k} f(T(k)) - \int_1^n \frac{1}{u} f(Z(u))\,\mathrm{d}u \right| &\leqslant C_3(\omega) + \int_1^n \frac{1}{u} |f(T(u)) - f(Z(u))|\,\mathrm{d}u \\
&\leqslant C_4(\omega) + \int_{K_\varepsilon^2}^n \frac{1}{u} |f(T(u)) - f(Z(u))|\,\mathrm{d}u
\end{aligned}$$

and it suffices to show

$$\frac{1}{(\log n)^{1/2-\eta}} \int_{K_\varepsilon^2}^n \frac{1}{u} |f(T(u)) - f(Z(u))|\,\mathrm{d}u \to 0 \quad \text{as } n \to \infty \qquad (16)$$

for almost all $\omega \in \Omega_\varepsilon$. By the triangle inequality it suffices to consider functions f satisfying either (3) or (4). We use the shorthand $\mathbb{E}_{(\varepsilon)}(X) = \mathbb{E}(XI(\Omega_\varepsilon)) = \int_{\Omega(\varepsilon)} X\,\mathrm{d}P$

for random variables X. Note that if $X \geqslant 0$ on Ω then $\mathbb{E}_{(\varepsilon)}(X) \leqslant \mathbb{E}(X)$. We will show that for some constants $u_0 > 0$ and $C_5 > 0$

$$\mathbb{E}_{(\varepsilon)} |f(T(u)) - f(Z(u))| \leqslant C_5(r_\varepsilon(u) + w(u)) \quad \text{for all } u \geqslant u_0 \tag{17}$$

where $w(u) = v(u)/u$. Since $r_\varepsilon(u)(\log u)^{1/2+2\eta} \to 0$ as $u \to \infty$ this will imply

$$\begin{aligned}
&\mathbb{E}_{(\varepsilon)}\left(\int_{K_\varepsilon^2}^{\infty} \frac{|f(T(u)) - f(Z(u))|}{u(\log u)^{1/2-\eta}}\,\mathrm{d}u\right) \\
&\quad \leqslant C_5 \int_{K_\varepsilon^2}^{\infty} \frac{r_\varepsilon(u)(\log u)^{1/2+2\eta}}{u(\log u)^{1+\eta}}\,\mathrm{d}u + C_5 \int_{K_\varepsilon^2}^{\infty} \frac{w(u)}{u(\log u)^{1/2-\eta}}\,\mathrm{d}u \\
&\quad \leqslant C_6 \int_{K_\varepsilon^2}^{\infty} \frac{\mathrm{d}u}{u(\log u)^{1+\eta}} + C_6 \int_{K_\varepsilon^2}^{\infty} \frac{\sqrt{\log^* u \log_3^* u}}{u^2(\log u)^{1/2-\eta}}\,\mathrm{d}u < \infty.
\end{aligned}$$

Hence, as $n \to \infty$, the sequence

$$\int_{K_\varepsilon^2}^{n} \frac{|f(T(u)) - f(Z(u))|}{u(\log u)^{1/2-\eta}}\,\mathrm{d}u = \mathcal{O}(1) + \sum_{k=[K_\varepsilon^2]+1}^{n} \int_{k-1}^{k} \frac{|f(T(u)) - f(Z(u))|}{u(\log u)^{1/2-\eta}}\,\mathrm{d}u$$

converges for almost every $\omega \in \Omega_\varepsilon$ to a finite number and Kronecker's Lemma (Breiman, 1992, Lemma 3.28) implies for $n \to \infty$

$$\begin{aligned}
0 &\leqslant \frac{1}{(\log n)^{1/2-\eta}} \int_{K_\varepsilon^2}^{n} \frac{1}{u}|f(T(u)) - f(Z(u))|\,\mathrm{d}u \\
&\leqslant \frac{1}{(\log n)^{1/2-\eta}} \sum_{k=[K_\varepsilon^2]+1}^{n} (\log k)^{1/2-\eta} \int_{k-1}^{k} \frac{|f(T(u)) - f(Z(u))|}{u(\log u)^{1/2-\eta}}\,\mathrm{d}u \to 0
\end{aligned}$$

which proves (16).

Now we are going to verify (17). We will consider three cases separately: First, f being an indicator, then f satisfying (3) and finally f satisfying (4). We define

$$Z_-(u) := Y_\Lambda(u - v(u)) - \log u \quad \text{and} \quad Z_+(u) := Y_\Lambda(u + v(u)) - \log u.$$

The form of the finite-dimensional distributions of Y_Λ (Resnick, 1987, p. 179, Proposition 4.7) and the identity $\Lambda(a - \log b) = \Lambda^b(a)$ for $a \in \mathbb{R}$, $b > 0$ give

$$P(Z_-(u) \leqslant s,\ Z_+(u) \leqslant t) = \Lambda^{1-w(u)}(\min\{s,t\})\Lambda^{2w(u)}(t) \tag{18}$$

and

$$P(Z(u) \leqslant z | Z_-(u) = s) = P(Z_+(u) \leqslant z | Z(u) = s) = \Lambda^{w(u)}(z) I(z \geqslant s) \tag{19}$$

for s, $z \in \mathbb{R}$ and $u \geqslant K_\varepsilon^2$. Note that the last conditional distribution function is absolutely continuous with density $w(u)\Lambda^{w(u)}(z)\mathrm{e}^{-z}I(z \geqslant s)$ except for a jump at $z = s$ of height $\Lambda^{w(u)}(s)$.

Suppose $f(x) = I(x > a)$ for some $a \in \mathbb{R}$. Then by (14), since f is non-decreasing,

$$\begin{aligned}
|f(T(u)) - f(Z(u))| &\leqslant f(Z_+(u) + r_\varepsilon(u)) - f(Z_-(u) - r_\varepsilon(u)) \\
&= \begin{cases} 1 & \text{if } Z_+(u) + r_\varepsilon(u) > a \text{ and } Z_-(u) - r_\varepsilon(u) \leqslant a, \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

Therefore, using (18),

$$
\begin{aligned}
&\mathbb{E}_{(\varepsilon)}(|f(T(u)) - f(Z(u))|)\\
&\quad\leqslant P(Z_-(u)\leqslant a + r_\varepsilon(u),\ Z_+(u) > a - r_\varepsilon(u))\\
&\quad= P(Z_-(u)\leqslant a + r_\varepsilon(u)) - P(Z_-(u)\leqslant a + r_\varepsilon(u),\ Z_+(u)\leqslant a - r_\varepsilon(u))\\
&\quad= \Lambda^{1-w(u)}(a + r_\varepsilon(u)) - \Lambda^{1+w(u)}(a - r_\varepsilon(u))\\
&\quad= \Lambda^{1-w(u)}(a + r_\varepsilon(u))\,(1 - \Lambda^{1+w(u)}(a - r_\varepsilon(u))/\Lambda^{1-w(u)}(a + r_\varepsilon(u)))\\
&\quad\leqslant \frac{\exp(a + r_\varepsilon(u))}{1 - w(u)}((1 + w(u))\mathrm{e}^{-a+r_\varepsilon(u)} - (1 - w(u))\mathrm{e}^{-a-r_\varepsilon(u)})
\end{aligned}
$$

by an application of the inequalities $\mathrm{e}^{-x}\leqslant 1/x$ and $1 - \mathrm{e}^{-x}\leqslant x$ for $x > 0$. Let u_0 be so large that $1 - w(u)\geqslant \frac{1}{2}$ and $r_\varepsilon(u)\leqslant 1$ for all $u\geqslant u_0$. Since $\mathrm{e}^{2x} - 1\leqslant \mathrm{e}^2 x$ for $x\in[0,1]$

$$
\begin{aligned}
\mathbb{E}_{(\varepsilon)}(|f(T(u)) - f(Z(u))|) &\leqslant 2((\mathrm{e}^{2r_\varepsilon(u)} - 1) + w(u)\,(\mathrm{e}^{2r_\varepsilon(u)} + 1))\\
&\leqslant 2(\mathrm{e}^2 + 1)\,(r_\varepsilon(u) + w(u))
\end{aligned}
$$

which is inequality (17). Note that this is a uniform bound in $a\in\mathbb{R}$.

Now let f be non-decreasing such that $\sup_{x\in\mathbb{R}}|f(x)| =: K < \infty$. Let again u_0 be so large that the estimates in the last paragraph hold. We will show that

$$
\mathbb{E}_{(\varepsilon)}|f(T(u)) - f(Z(u))|\leqslant 4K(\mathrm{e}^2 + 1)\,(r_\varepsilon(u) + w(u)) \tag{20}
$$

for all $u\geqslant u_0$. Fix $u\geqslant u_0$. Since f is non-decreasing

$$
|f(T(u)) - f(Z(u))|\leqslant f(Z_+(u) + r_\varepsilon(u)) - f(Z_-(u) - r_\varepsilon(u))
$$

and there is a sequence $\{f_m,\ m\in\mathbb{N}\}$ of non-decreasing step functions converging almost everywhere to $f(x)$ and $\sup_{x\in\mathbb{R}}|f_m(x)|\leqslant K$ for all $m\in\mathbb{N}$. For each function f_m there are $n(m)\in\mathbb{N}$, $a_1,\dots,a_{n(m)}\in\mathbb{R}$ and numbers $b_0,\dots,b_{n(m)}$ with $b_j > 0$, $j = 1,\dots,n(m)$ such that

$$
f_m(x) = b_0 + \sum_{j=1}^{n(m)} b_j I(x > a_j) \quad\text{and}\quad \sum_{j=1}^{n(m)} b_j\leqslant 2K.
$$

Applying the bounded convergence theorem, we see that given $\delta > 0$, there is a $m_0\in\mathbb{N}$ such that

$$
\mathbb{E}|f(Z_+(u) + r_\varepsilon(u)) - f_m(Z_+(u) + r_\varepsilon(u))|\leqslant\delta
$$

and

$$
\mathbb{E}|f(Z_-(u) - r_\varepsilon(u)) - f_m(Z_-(u) - r_\varepsilon(u))|\leqslant\delta
$$

for all $m\geqslant m_0$. Hence for $m\geqslant m_0$

$$
\begin{aligned}
&\mathbb{E}_{(\varepsilon)}(|f(T(u)) - f(Z(u))|)\\
&\quad\leqslant \mathbb{E}|f(Z_+(u) + r_\varepsilon(u)) - f_m(Z_+(u) + r_\varepsilon(u))|\\
&\qquad+ \mathbb{E}_{(\varepsilon)}|f_m(Z_+(u) + r_\varepsilon(u)) - f_m(Z_-(u) - r_\varepsilon(u))|
\end{aligned}
$$

$$+\,\mathbb{E}|f_m(Z_-(u)-r_\varepsilon(u))-f(Z_-(u)-r_\varepsilon(u))|$$

$$\leqslant 2\delta+\sum_{j=1}^{n(m)} b_j\mathbb{E}_{(\varepsilon)}(I(Z_+(u)+r_\varepsilon(u)>a_j)-I(Z_-(u)-r_\varepsilon(u)>a_j))$$

$$\leqslant 2\delta+\sum_{j=1}^{n(m)} b_j 2(\mathrm{e}^2+1)(r_\varepsilon(u)+w(u))$$

$$\leqslant 2\delta+4K(\mathrm{e}^2+1)(r_\varepsilon(u)+w(u)).$$

Since $\delta>0$ is arbitrary, (20) follows.

Finally, suppose f satisfies (4). Elementary calculus yields the finiteness of the following integrals for $0<\gamma<\frac{1}{2}$:

$$\int_{-\infty}^{\infty} h_\gamma(z)\,\mathrm{d}\Lambda(z),\quad \int_{-\infty}^{\infty} h_\gamma(z)\mathrm{e}^{-z}\,\mathrm{d}\Lambda(z),\quad \int_{-\infty}^{\infty} h_\gamma(z-1)\,\mathrm{d}\Lambda(z), \tag{21}$$

$$\int_{-\infty}^{\infty} h_\gamma(s-1)\mathrm{e}^{-2s}\Lambda^{1/2}(s)\,\mathrm{d}s,\quad \int_{-\infty}^{\infty} \mathrm{e}^{(\gamma-1)s}\Lambda^{1/2}(s)\mathrm{e}^{-s}\,\mathrm{d}s. \tag{22}$$

For the rest of the proof we will drop the arguments of our processes, writing T instead of $T(u)$ and similarly for the other processes, whenever this is convenient. We have

$$\mathbb{E}_{(\varepsilon)}|f(T)-f(Z)|$$

$$=\mathbb{E}_{(\varepsilon)}(|f(T)-f(Z)|I(T>Z))+\mathbb{E}_{(\varepsilon)}(|f(T)-f(Z)|I(T\leqslant Z))$$

$$=\mathbb{E}_{(\varepsilon)}(|f(Z+D_1)-f(Z)|I(T>Z))$$

$$+\,\mathbb{E}_{(\varepsilon)}(|f(T+D_2)-f(T)|I(T\leqslant Z)),$$

where by (14) for all $u\geqslant K_\varepsilon^2$ and almost all $\omega\in\Omega_\varepsilon$

$$0\leqslant D_1:=T-Z\leqslant Z_+-Z+r_\varepsilon \quad\text{and}\quad 0\leqslant D_2:=Z-T\leqslant Z-Z_-+r_\varepsilon.$$

Thus by (4) and monotonicity

$$\mathbb{E}_{(\varepsilon)}|f(T)-f(Z)|$$

$$\leqslant C_7\mathbb{E}_{(\varepsilon)}(h_\gamma(Z)(\mathrm{e}^{\gamma D_1}-1))+C_7\mathbb{E}_{(\varepsilon)}(h_\gamma(T)(\mathrm{e}^{\gamma D_2}-1)I(T\leqslant Z))$$

$$\leqslant C_7\mathbb{E}(h_\gamma(Z)(\mathrm{e}^{\gamma(Z_+-Z+r_\varepsilon)}-1))$$

$$+\,C_7\mathbb{E}((\mathrm{e}^{\gamma Z}+\exp(\mathrm{e}^{-\gamma(Z_--r_\varepsilon)})\mathrm{e}^{-\gamma(Z_--r_\varepsilon)})(\mathrm{e}^{\gamma(Z-Z_-+r_\varepsilon)}-1))$$

$$=C_7(I+II).$$

Let $u_0>0$, such that $r_\varepsilon(u)\leqslant 1$ and $w(u)\leqslant\frac{1}{2}$ for all $u\geqslant u_0$. Using well-known properties of conditional expectation (cf. Definition 4.12 and Theorem 4.28 of Breiman, 1992) and (19) we get for $u \geqslant u_0$

$$
\begin{aligned}
I &= \mathbb{E}(\mathbb{E}(h_\gamma(Z)\,(\mathrm{e}^{\gamma(Z_+ - Z + r_\varepsilon)} - 1)|Z = z)) \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} h_\gamma(z)\,(\mathrm{e}^{\gamma(t-z+r_\varepsilon)} - 1)\,\mathrm{d}P(Z_+ \leqslant t|Z = z)\,\mathrm{d}P(Z \leqslant z) \\
&= \int_{-\infty}^{\infty} h_\gamma(z)\,(\mathrm{e}^{\gamma r_\varepsilon} - 1)\Lambda^{w(u)}(z)\,\mathrm{d}\Lambda(z) \\
&\quad + \int_{-\infty}^{\infty} h_\gamma(z) \int_z^{\infty} (\mathrm{e}^{\gamma(t-z+r_\varepsilon)} - 1) w(u) \Lambda^{w(u)}(t)\mathrm{e}^{-t}\,\mathrm{d}t\,\mathrm{d}\Lambda(z) \\
&\leqslant C_8 r_\varepsilon(u) \int_{-\infty}^{\infty} h_\gamma(z)\,\mathrm{d}\Lambda(z) + w(u)\mathrm{e}^{\gamma} \int_{-\infty}^{\infty} h_\gamma(z)\mathrm{e}^{-\gamma z} \int_z^{\infty} \mathrm{e}^{(\gamma-1)t}\,\mathrm{d}t\,\mathrm{d}\Lambda(z) \\
&\leqslant C_9 r_\varepsilon(u) + C_{10} w(u) \int_{-\infty}^{\infty} h_\gamma(z)\mathrm{e}^{-z}\,\mathrm{d}\Lambda(z) \leqslant C_{11}(r_\varepsilon(u) + w(u))
\end{aligned}
$$

by (21). Similarly, we get for the second term for $u \geqslant u_0$

$$
\begin{aligned}
II &\leqslant \mathbb{E}\,(\mathrm{e}^{\gamma Z} + h_\gamma(Z_- - 1))\,(\mathrm{e}^{\gamma(Z - Z_- + r_\varepsilon)} - 1) \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} (\mathrm{e}^{\gamma z} + h_\gamma(s-1))\,(\mathrm{e}^{\gamma(z-s+r_\varepsilon)} - 1)\,\mathrm{d}P(Z \leqslant z|Z_- = s)\,\mathrm{d}P(Z_- \leqslant s) \\
&\leqslant \int_{-\infty}^{\infty} (h_\gamma(s) + h_\gamma(s-1))\,(\mathrm{e}^{\gamma r_\varepsilon} - 1)\Lambda^{w(u)}(s)\Lambda^{1-w(u)}(s)\mathrm{e}^{-s}\,\mathrm{d}s \\
&\quad + w(u)\mathrm{e}^{\gamma} \int_{-\infty}^{\infty}\int_s^{\infty} (\mathrm{e}^{\gamma z} + h_\gamma(s-1))\mathrm{e}^{\gamma(z-s)}\Lambda^{w(u)}(z)\mathrm{e}^{-z}\,\mathrm{d}z\,\mathrm{d}P(Z_- \leqslant s) \\
&\leqslant C_{12} r_\varepsilon \int_{-\infty}^{\infty} (h_\gamma(s) + h_\gamma(s-1))\,\mathrm{d}\Lambda(s) + C_{13} w(u) \\
&\leqslant C_{14}(r_\varepsilon(u) + w(u)).
\end{aligned}
\tag{23}
$$

The inner integral in (23) can be bounded by

$$
\begin{aligned}
&\mathrm{e}^{-\gamma s} \int_s^{\infty} \mathrm{e}^{(2\gamma-1)z}\,\mathrm{d}z + h_\gamma(s-1)\mathrm{e}^{-\gamma s} \int_s^{\infty} \mathrm{e}^{\gamma z}\mathrm{e}^{-z}\,\mathrm{d}z \\
&\quad \leqslant C_{15}(\mathrm{e}^{(\gamma-1)s} + h_\gamma(s-1)\mathrm{e}^{-s}).
\end{aligned}
$$

Thus, since $w(u) \leqslant \frac{1}{2}$ for $u \geqslant u_0$, an upper estimate for the double integral of (23) is

$$
C_{16} \int_{-\infty}^{\infty} (\mathrm{e}^{(\gamma-1)s} + h_\gamma(s-1)\mathrm{e}^{-s})\Lambda^{1/2}(s)\mathrm{e}^{-s}\,\mathrm{d}s
$$

which is finite by (22). This completes the proof of Theorem 1. □

Proof of Theorem 2. The proof follows standard arguments introduced by Dœblin (1938). See also Mandl (1968, p. 92ff), Motoo (1959), Itô and McKean (1965, p. 228f), Robbins and Siegmund (1971) and Horváth and Khoshnevisan (1996). We will exploit the Markovian structure of the process U. Among the many good references on Markov processes we choose Blumenthal and Getoor (1968) as our main reference.

Let $\mathscr{F}_t^0 := \sigma(U(s), s \leqslant t)$ for $t \in \mathbb{R}$ and $\mathscr{F}^0 := \sigma(U(s), s \in \mathbb{R})$ and for $x \in \mathbb{R}$ we denote by P^x the probability measure $P^x(A) := P(A \mid U(0) = x)$ for $A \in \mathscr{F}^0$. Given a finite measure μ on the Borel sets $\mathscr{B}(\mathbb{R})$ of $\mathbb{R}$, define the finite measure P^μ (Blumenthal and Getoor, 1968, p. 25) by

$$P^\mu(A) = \int_{\mathbb{R}} P^x(A)\,\mathrm{d}\mu(x) \quad \text{for } A \in \mathscr{F}^0. \tag{24}$$

$\mathbb{E}^x$ resp. $\mathbb{E}^\mu$ denotes the expectation operator with respect to the measures P^x resp. P^μ. Moreover, $\mathscr{F}$ denotes the completion of $\mathscr{F}^0$ with respect to the family of finite measures $\{P^\mu; \mu \text{ a finite measure on } \mathscr{B}(\mathbb{R})\}$ and let $\mathscr{F}_t$ be the completion of $\mathscr{F}_t^0$ in $\mathscr{F}$ with respect to the same family (cf. Blumenthal and Getoor, 1968, p. 27). Finally,

$$\mathscr{F}_{t+} := \bigcap_{s>t} \mathscr{F}_s$$

and $\theta(t)$ denotes the shift operator defined by $U(s) \circ \theta(t) = U(s+t)$ for $s,\ t \in \mathbb{R}$. Then U has the following properties:

Lemma 7. (a) $(\Omega, \mathscr{F}, \mathscr{F}_{t+}, U(t), \theta(t), P^x)$ *is a temporally homogeneous right continuous strong Markov process, i.e. for every bounded $\mathscr{F}$-measurable random variable Y we have the strong Markov property*

$$\mathbb{E}^x(Y \circ \theta(T) \mid \mathscr{F}_T) = \mathbb{E}^{U(T)}(Y) \tag{25}$$

for all stopping times T of the filtration $\{\mathscr{F}_{t+},\ t > 0\}$ and all $x \in \mathbb{R}$. Under P^Λ, where $\Lambda(A) := \int_A \mathrm{d}\Lambda(t)$ for $A \in \mathscr{B}(\mathbb{R})$, $U(t)$ is stationary and $P(U(s) \leqslant u) = P^\Lambda(U(s) \leqslant u) = \Lambda(u)$ for all $u, s \in \mathbb{R}$.

(b) *For P-almost all $\omega \in \Omega$, U crosses the interval $[-1, 1]$ infinitely often.*

(c) *There exist $K_1,\ K_2 > 0$ such that $|\gamma_U(r)| \leqslant K_1 \mathrm{e}^{-K_2|r|}$, $r \in \mathbb{R}$, where $\gamma_U(r) := \mathrm{Cov}(f(U_r), f(U_0))$.*

Remark 8. Using the definition of conditional expectation and definition (24), it can easily be seen that if $U(T) \equiv 0$ the strong Markov property (25) implies

$$\mathbb{E}^\mu(Y \circ \theta(T) I(A)) = \mathbb{E}^0(Y) P^\mu(A) \quad \text{for all } A \in \mathscr{F}_T \tag{26}$$

where μ is a finite measure.

Proof of Lemma 7. (a) By construction, Y_Λ is a right continuous Markov jump process, thus U is also Markovian and right continuous and we have

$$\begin{aligned} P(U(r+s) \leqslant x | U(s) = y) &= \Lambda^{\mathrm{e}^r - 1}(x+r) I(x + r \geqslant y) \\ &= \Lambda^{1-\mathrm{e}^{-r}}(x) I(x \geqslant y - r) \end{aligned} \tag{27}$$

for $r, s > 0$, $x, y \in \mathbb{R}$. This probability does not depend on s; thus, U is temporally homogeneous. The strong Markov property can be established similarly to the calculation in Blumenthal and Getoor (1968, p. 67), where it is shown that U is in fact a standard process. The key step is Yushkevitch's theorem (Blumenthal and Getoor, 1968, Theorem 8.11). Relation (25) is Corollary 8.6 of Blumenthal and Getoor (1968). The last claim follows from (24) and the form of the finite-dimensional distributions of Y_Λ.

(b) By definition of U the sample paths are right continuous with jumps of positive jump heights. Between the jumps the path consists of straight lines with slope -1. Thus it suffices to show

$$P(U(\log n) > 1 \text{ i.o.}) = P(U(\log n) < -1 \text{ i.o.}) = 1.$$

Note that $\{Y_\Lambda(n),\ n\in\mathbb{N}\} \overset{\mathrm{d}}{=} \{M_n',\ n\in\mathbb{N}\}$ in $\mathbb{R}^\infty$ where $M_n' = \max_{1\leqslant k\leqslant n} X_k'$ with a sequence of i.i.d. random variables X_k' with common distribution function Λ (Dwass, 1964). Thus,

$$P(U(\log n) > 1 \text{ i.o.}) = P(M_n' > \log n + 1 \text{ i.o.}) = P(X_n' > \log n + 1 \text{ i.o.}).$$

By the Borel–Cantelli lemma, since

$$\sum_{n\in\mathbb{N}} P(X_n' > \log n + 1) = \sum_{n\in\mathbb{N}} \bar{\Lambda}(\log n + 1) \sim \frac{1}{e}\sum_{n\in\mathbb{N}} \frac{1}{n} = \infty,$$

this probability is 1. Also,

$$P(U(\log n) < -1 \text{ i.o.}) = P(M_n' < \log n - 1 \text{ i.o.}) = 1,$$

since $n\bar{\Lambda}(\log n - 1) \sim e^{-1} \nrightarrow \infty$, so we can apply Theorem 3.5.2 of Embrechts et al. (1997).

(c) Obviously, $|\gamma_U(r)| \leqslant \mathbb{E}^\Lambda f(U_0)^2 < \infty$ and by symmetry it suffices to consider $r \geqslant r_0$ for some $r_0 > 0$. We have by (27)

$$\begin{aligned}
\gamma_U(r) &= \int_{-\infty}^{\infty} f(y) \int_{-\infty}^{\infty} f(x)\,\mathrm{d}P(U_r \leqslant x | U_0 = y)\,\mathrm{d}P(U_0 \leqslant y) - m^2 \\
&= \int_{-\infty}^{\infty} f(y) f(y-r) \Lambda^{\mathrm{e}^{r-1}}(y)\,\mathrm{d}\Lambda(y) \\
&\quad + \int_{-\infty}^{\infty}\int_{y-r}^{\infty} f(x)f(y)(1-\mathrm{e}^{-r})\Lambda^{1-\mathrm{e}^{-r}}(x)\mathrm{e}^{-x}\,\mathrm{d}x\,\mathrm{d}\Lambda(y) - m^2 \\
&= \int_{-\infty}^{\infty} f(y) f(y-r) \Lambda^{\mathrm{e}^{r}-1}(y)\,\mathrm{d}\Lambda(y) \\
&\quad + \int_{-\infty}^{\infty}\int_{y-r}^{\infty} f(x)f(y)((1-\mathrm{e}^{-r})\Lambda^{1-\mathrm{e}^{-r}}(x) - \Lambda(x))\mathrm{e}^{-x}\,\mathrm{d}x\,\mathrm{d}\Lambda(y) \\
&\quad - \int_{-\infty}^{\infty}\int_{-\infty}^{y-r} f(x)f(y)\,\mathrm{d}\Lambda(x)\,\mathrm{d}\Lambda(y).
\end{aligned}$$

Recall that $|f(x)| \leqslant C_1 h_\gamma(x)$ for some $C_1 > 0,\ 0 < \gamma < \frac{1}{2}$, thus

$$|\gamma_U(r)| \leqslant C_1^2 \left(\int_{-\infty}^{\infty} h_\gamma(y) h_\gamma(y-r) \Lambda^{e^r - 1}(y) \, d\Lambda(y) \right.$$
$$+ \int_{-\infty}^{\infty} h_\gamma(y) \int_{y-r}^{\infty} h_\gamma(x) |(1-e^{-r})\Lambda^{1-e^{-r}}(x) - \Lambda(x)| e^{-x} \, dx \, d\Lambda(y)$$
$$\left. + \int_{-\infty}^{\infty} h_\gamma(y) \int_{-\infty}^{y-r} h_\gamma(x) \, d\Lambda(x) \, d\Lambda(y) \right) = C_1^2 (I + II + III).$$

Using the definition of $h_\gamma(y-r)$ and substituting $w = e^{-(y-r)}$ with respect to y we get

$$I \leqslant e^{-\gamma r} \int_{-\infty}^{\infty} h_\gamma(y) e^{\gamma y} \, d\Lambda(y) + \int_{-\infty}^{\infty} h_\gamma(y) \exp(e^{-\gamma(y-r)}) e^{-\gamma(y-r)} \Lambda^{e^r - 1}(y) \, d\Lambda(y)$$
$$\leqslant e^{-\gamma r} \int_{-\infty}^{\infty} h_\gamma^2(y) \, d\Lambda(y) + \int_0^{\infty} (w^{-\gamma} e^{\gamma r} + e^{w^\gamma} w^\gamma) e^{w^\gamma} w^\gamma e^{-w} e^{-r} \, dw$$
$$\leqslant C_2 e^{-\gamma r} + e^{-(1-\gamma) r} \int_0^{\infty} e^{w^\gamma - w} \, dw + e^{-r} \int_0^{\infty} w^{2\gamma} e^{2w^\gamma - w} \, dw$$
$$\leqslant C_3 e^{-\gamma r}.$$

For the second term, apply the inequality $|tz^t - z| \leqslant z^t (1 - \log z)(1-t)$ for $0 < t,\ z \leqslant 1$ to $z = \Lambda(x)$ and $t = 1 - e^{-r}$. Let r_0 be so large that $1 - e^{-r} \geqslant \frac{1}{2}$ for $r \geqslant r_0$. Whence for $r \geqslant r_0$

$$II \leqslant \int_{-\infty}^{\infty} h_\gamma(y) \, d\Lambda(y) \int_{-\infty}^{\infty} h_\gamma(x) \Lambda^{1/2}(x) (1 + e^{-x}) e^{-x} \, dx e^{-r} \leqslant C_4 e^{-r}.$$

For the third term, note that the substitution $w = e^{-x}$ and the inequality $e^{-x} \leqslant 1/x$ for $x > 0$ yield

$$\int_{-\infty}^{y-r} h_\gamma(x) \, d\Lambda(x) = \int_{e^{r-y}}^{\infty} w^{-\gamma} e^{-w} \, dw + \int_{e^{r-y}}^{\infty} e^{-w^\gamma} w^\gamma e^{2w^\gamma - w} \, dw$$
$$\leqslant (e^{r-y})^{-\gamma} \int_0^{\infty} e^{-w} \, dw + \exp(-e^{\gamma(r-y)}) \int_0^{\infty} w^\gamma e^{2w^\gamma - w} \, dw$$
$$\leqslant e^{-\gamma r} (h_\gamma(y) + C_5 h_\gamma(y)),$$

thus $III \leqslant C_6 e^{-\gamma r}$. $\square$

Fix $f \in \mathscr{M}_\varrho,\ 0 < \varrho < \frac{1}{4}$. By part (b) of the previous lemma, the following quantities are P^Λ-a. s. well-defined for $k \in \mathbb{N}$:

$$\tau_0 := 0,$$
$$\tau_{2k-1} := \inf\{s:\ s > \tau_{2k-2},\ U(s) = 1\},$$
$$\tau_{2k} := \inf\{s:\ s > \tau_{2k-1},\ U(s) = 0\}$$

and

$$\xi_k := \int_{\tau_{2k}}^{\tau_{2k+2}} (f(U_s) - m)\,\mathrm{d}s.$$

Then we have

Lemma 9. (a) $\{\tau_k,\ k \in \mathbb{N}\}$ *is a sequence of* $\mathscr{F}_{t+}$*-stopping times and* $\{\tau_{2k+2} - \tau_{2k},\ k \in \mathbb{N}\}$ *forms a sequence of i.i.d. random variables with respect to* P^Λ.
(b) $\{\xi_k,\ k \in \mathbb{N}\}$ *is a sequence of i.i.d. random variables with respect to* P^Λ.
(c) *We have* $\mathbb{E}^\Lambda(\tau_2^k) < \infty$ *and* $\mathbb{E}^\Lambda(\tau_4 - \tau_2)^k < \infty$ *for all* $k \in \mathbb{N}$.
(d) *For all* $k \in \mathbb{N}$ *there is a constant* $K(k)$ *such that* $P(\tau_2 \leqslant s \leqslant \tau_4) \leqslant \min\{K(k)/s^k, 1\}$.
(e) $\mathbb{E}^\Lambda|\xi_1|^{2/(1-2\varrho)} \leqslant \mathbb{E}^\Lambda(\int_{\tau(2)}^{\tau(4)} |f(U_s) - m|\,\mathrm{d}s)^{2/(1-2\varrho)} < \infty$.
(f) $\mathbb{E}^\Lambda \xi_1 = 0$ *and* $\mathbb{E}^\Lambda(\xi_1^2) = \mu \int_{-\infty}^{\infty} \gamma_U(|r|)\,\mathrm{d}r = \mu\sigma^2$ *where* $\mu = \mathbb{E}^\Lambda(\tau_4 - \tau_2)$ *and* σ^2 *is given by* (8) *in Theorem* 2.

Proof of Lemma 9. (a) The times τ_k are successive hitting times of the Borel sets $\{0\}$ resp. $\{1\}$. Thus, Theorem 10.7 of Blumenthal and Getoor (1968) implies that τ_k are $\mathscr{F}_{t+}$-stopping times. Let $A,\ B \in \mathscr{B}(\mathbb{R})$ and set $Y := I(\tau_6 - \tau_4 \in A)$. Note that $Y \circ \theta(\tau_4) = Y$ and $\{\tau_4 - \tau_2 \in B\} \in \mathscr{F}_{\tau(4)}$, hence by (26)

$$P^\Lambda(\tau_6 - \tau_4 \in A, \tau_4 - \tau_2 \in B) = P^\Lambda(\tau_4 - \tau_2 \in B) P^0(\tau_6 - \tau_4 \in A).$$

Setting $B = \mathbb{R}$ yields $P^0(\tau_6 - \tau_4 \in A) = P^\Lambda(\tau_6 - \tau_4 \in A)$, thus $\tau_4 - \tau_2$ is independent of $\tau_6 - \tau_4$ with respect to P^Λ. Next recall from Proposition 10.2 of Blumenthal and Getoor (1968) that $\tau_{2k+2} = \tau_{2k} + \tau_2 \circ \theta(\tau_{2k})$ for $k \in \mathbb{N}$. Let $Y := I(\tau_2 \in A)$, then $Y \circ \theta(\tau_2) = I(\tau_4 - \tau_2 \in A)$ and $Y \circ \theta(\tau_4) = I(\tau_6 - \tau_4 \in A)$, thus by (26) $\tau_4 - \tau_2$ and $\tau_6 - \tau_4$ have identical distributions. The general case is proved similarly. More details can be found in Fahrner (2000b).
(b) Again let $A \in \mathscr{B}(\mathbb{R})$, apply (26) to $Y := I(\xi_2 \in A)$ and observe that $Y \circ \theta(\tau_4) = Y$, hence we have independence between ξ_1 and ξ_2. It can be seen easily that these random variables also have identical distributions, let $Y := I(\int_0^{\tau(2)} (f(U_s) - m)\,\mathrm{d}s \in A)$ and note that $Y \circ \theta(\tau_2) = I(\xi_1 \in A)$ and $Y \circ \theta(\tau_4) = I(\xi_2 \in A)$.
(c) Define a process V with time space $\mathbb{N}_0$ and state space $\mathbb{Z}$ by $V(n) = [U(n)],\ n \in \mathbb{N}_0$, where $[x]$ denotes the greatest integer less than or equal to x, i.e., $V(n) = k$ iff $U(n) \in [k, k+1)$. By writing down the definition of a Markov process, we see that passing to a discrete-time skeleton preserves the Markovian structure. Thus, V is a Markov chain and since U is stationary under P^Λ, so is V. This stationary distribution assigns positive probability $\Lambda(k+1) - \Lambda(k)$ to state k, thus each state is recurrent (Durrett, 1996, Theorem 5.4.5). The connection to our problem is the following: If V visits state 1 and then state -1 before time n then U hits 1 and 0 before n, therefore let $T_x := \inf\{n > 0\colon V(n) = x\}$ denote the hitting time of $x \in \mathbb{Z}$ for V, then for $k \in \mathbb{N}$:

$$\mathbb{E}^x \tau_2^k \leqslant 2^k(\mathbb{E}^{[x]} T_1^k + \mathbb{E}^1 T_{-1}^k) \quad \text{and} \quad \mathbb{E}^x(\tau_4 - \tau_2)^k \leqslant 2^k(\mathbb{E}^0 T_1^k + \mathbb{E}^1 T_{-1}^k).$$

Whence it suffices to consider moments of T_x and since each state of V is recurrent, $\mathbb{E}^x T_x^k < \infty$ implies $\mathbb{E}^y T_x^k < \infty$ for all $x, y \in \mathbb{Z}$, $k \in \mathbb{N}$, i.e., the starting point of the process is inessential.

This can be seen as follows: Let $n := \min\{m \in \mathbb{N}: P^x(V_m = y) > 0\}$. Choose $y_1, \dots, y_{n-1} \neq y$ with $P^x(A) > 0$ where $A := \{V_1 = y_1, \dots, V_{n-1} = y_{n-1}, V_n = y\} \in \mathcal{G}_n := \sigma(V_1, \dots, V_n)$. Then

$$\mathbb{E}^x T_x^k \geqslant \int_A T_x^k \, \mathrm{d}P^x = \int_A \mathbb{E}^x((T_x - n)^k \circ \theta(n) | \mathcal{G}_n) \, \mathrm{d}P^x$$

note $T_x - n \geqslant 0$ a.s. by definition of A and by the Markov property

$$= \int_A \mathbb{E}^{V(n)}(T_x - n)^k \, \mathrm{d}P^x = \mathbb{E}^y (T_x - n)^k P^x(A),$$

i.e., $\mathbb{E}^x T_x^k < \infty$ implies $\mathbb{E}^y (T_x - n)^k < \infty$ which implies $\mathbb{E}^y T_x^k < \infty$.

We will also use the following first entrance decomposition which is an easy consequence of the Markov property of V:

$$P^x(V(n) = y) = \sum_{m=1}^{n} P^x(T_y = m) P^y(V(n-m) = y). \tag{28}$$

Set $v_n := P^y(V(n) = y)$, $f_m := P^y(T_y = m)$, $b_0 = 1$ and $b_n = 0$ for $n \geqslant 1$, then (28) reads

$$v_n = b_n + \sum_{m=1}^{n} f_m v_{n-m}$$

and passing to generating functions (denoted by the corresponding capital letters with a hat) we get (Feller, 1968, p. 330)

$$\hat{V}(s) = \frac{\hat{B}(s)}{1 - \hat{F}(s)} = \frac{1}{1 - \hat{F}(s)} \Leftrightarrow \hat{F}(s) = 1 - \frac{1}{\hat{V}(s)}. \tag{29}$$

It is well known that if the generating function $\hat{F}$ has a finite nth derivative at 1 then $\mathbb{E}^y(T_y^n) < \infty$. We have by Definition 4.7 of Breiman (1992)

$$\begin{aligned}
v_n &= P(V(k+n) = y | V(k) = y) \\
&= P(U(k+n) \in [y, y+1) | U(k) \in [y, y+1)) \\
&= \int_y^{y+1} P(U(k+n) \in [y, y+1) | U(k) = t) \, \mathrm{d}\Lambda(t) / P(U(k) \in [y, y+1)) \\
&= \begin{cases} 1 & \text{if } n = 0, \\ \Lambda^{1-\mathrm{e}^{-n}}(y+1) - \Lambda^{1-\mathrm{e}^{-n}}(y) & \text{if } n \geqslant 1, \end{cases}
\end{aligned}$$

such that (with $g := \Lambda(y+1) - \Lambda(y)$)

$$\begin{aligned}
\hat{V}(s) &= 1 + \sum_{n=1}^{\infty} (\Lambda^{1-\mathrm{e}^{-n}}(y+1) - \Lambda^{1-\mathrm{e}^{-n}}(y)) s^n \\
&= \sum_{n=0}^{\infty} g s^n + 1 - g + \Lambda(y+1) \sum_{n=1}^{\infty} (\exp(\mathrm{e}^{-n-y-1}) - 1) s^n \\
&\quad - \Lambda(y) \sum_{n=1}^{\infty} (\exp(\mathrm{e}^{-n-y}) - 1) s^n \\
&=: \frac{g}{1-s} + H(s).
\end{aligned}$$

Since $(\exp(-e^{-n-y})-1)s^n \sim -e^{-y}(s/e)^n$, we see that H is a holomorphic function on $|s|<e$. Therefore by (29)

$$\hat{F}'(s)=\frac{\hat{V}'(s)}{\hat{V}^2(s)}=\frac{H_1(s)}{g^2+(1-s)G_1(s)},$$

where $H_1(s)=g+H'(s)(1-s)^2$ and $G_1(s)=2H(s)+(1-s)$ are holomorphic functions on $|s|<e$. Now an induction shows for the nth derivative:

$$\hat{F}^{(n)}(s)=\frac{H_n(s)}{g^{2^n}+(1-s)G_n(s)}$$

with holomorphic functions $H_n(s)$ and $G_n(s)$ on $|s|<e$. Thus, $\lim_{s\uparrow 1}\hat{F}^{(n)}(s)$ exists for all $n\in\mathbb{N}$ and the claim follows.
(d) This follows from (c) and the Markov inequality:

$$P^{\Lambda}(\tau_2\leqslant s\leqslant\tau_4)\leqslant P^{\Lambda}(\tau_4\geqslant s)\leqslant s^{-k}\mathbb{E}^{\Lambda}(\tau_4^k)\leqslant(2/s)^k(\mathbb{E}^{\Lambda}(\tau_2^k)+\mathbb{E}^{\Lambda}(\tau_4-\tau_2)^k)$$

for all $k\in\mathbb{N}$.
(e) Recall (11), i.e., $|f(x)|\leqslant Ch_\gamma(x)$ with $\gamma/(1/2-\varrho)<1$. Choose $p_1,p_2>1$ such that $\gamma p_1 p_2/(\frac{1}{2}-\varrho)<1$. Let $p_3=p_1/(\frac{1}{2}-\varrho)$ and q_i denotes the conjugate of p_i, i.e., $1/p_i+1/q_i=1,\ i=1,2,3$. By Hölder's inequality we have

$$\begin{aligned}|\xi_1| &\leqslant \int_{\tau(2)}^{\tau(4)}|f(U_s)-m|\,\mathrm{d}s=\int_{\mathbb{R}}|f(U_s)-m|I(\tau_2\leqslant s\leqslant\tau_4)I(\tau_2\leqslant s\leqslant\tau_4)\,\mathrm{d}s\\ &\leqslant\left(\int_{\mathbb{R}}|f(U_s)-m|^{p_3}I(\tau_2\leqslant s\leqslant\tau_4)\,\mathrm{d}s\right)^{1/p_3}\left(\int_{\mathbb{R}}I(\tau_2\leqslant s\leqslant\tau_4)\,\mathrm{d}s\right)^{1/q_3}.\end{aligned}$$

Thus,

$$\begin{aligned}|\xi_1|^{2/(1-2\varrho)} &\leqslant\left(\int_{\tau(2)}^{\tau(4)}|f(U_s)-m|\,\mathrm{d}s\right)^{2/(1-2\varrho)}\\ &\leqslant\left(\int_{\tau(2)}^{\tau(4)}|f(U_s)-m|^{p_3}\,\mathrm{d}s\right)^{1/p_1}(\tau_4-\tau_2)^q\end{aligned}$$

where $q=2/((1-2\varrho)q_3)$. Taking expectations and applying Hölder's inequality with $p=p_1$ yields

$$\begin{aligned}&\mathbb{E}^{\Lambda}\left(\int_{\tau(2)}^{\tau(4)}|f(U_s)-m|\,\mathrm{d}s\right)^{2/(1-2\varrho)}\\ &\quad\leqslant\left(\mathbb{E}^{\Lambda}\left(\int_{\tau(2)}^{\tau(4)}|f(U_s)-m|^{p_3}\,\mathrm{d}s\right)\right)^{1/p_1}(\mathbb{E}^{\Lambda}(\tau_4-\tau_2)^{q\cdot q_1})^{1/q_1}\\ &\quad\leqslant\left(\int_{\mathbb{R}}\mathbb{E}^{\Lambda}(|f(U_s)-m|^{p_3}I(\tau_2\leqslant s\leqslant\tau_4))\,\mathrm{d}s\right)^{1/p_1}C\end{aligned}$$

and applying Hölder's inequality with $p = p_2$ this can be bounded by

$$\leqslant C\left(\int_{\mathbb{R}} (\mathbb{E}^{\Lambda}|f(U_s) - m|^{p_2 \cdot p_3})^{1/p_2}(P(\tau_2 \leqslant s \leqslant \tau_4))^{1/q_2}\, \mathrm{d}s\right)^{1/p_1}$$

$$\leqslant C(\mathbb{E}^{\Lambda}|f(U_0) - m|^{p_2 \cdot p_3})^{1/p_1 p_2}\left(\int_{\mathbb{R}} P(\tau_2 \leqslant s \leqslant \tau_4)^{1/q_2}\, \mathrm{d}s\right)^{1/p_1} < \infty,$$

since $(f(x))^{p_2 p_3} \leqslant C' h_{\gamma'}(x)$ with some $\gamma' < 1$ and we have the estimate of part (c).
(f) For $k \in \mathbb{N}$, $s, t \in \mathbb{R}$, $A_1, A_2, \ldots, A_k$, $B_1, B_2 \in \mathscr{B}(\mathbb{R})$ letting $Y = I(U(s - \tau_{2k}) \in B_1, U(t - \tau_{2k}) \in B_2) \in \mathscr{F}$, $T = \tau_{2k}$ and $A = \{\tau_2 \in A_1, \tau_4 \in A_2, \ldots, \tau_{2k} \in A_k\} \in \mathscr{F}_{\tau(2k)}$ in (26) we get

$$P^{\Lambda}(U_s \in B_1, U_t \in B_2, \tau_2 \in A_1, \ldots, \tau_{2k} \in A_k)$$
$$= P^0(U(s - \tau_{2k}) \in B_1,\ U(t - \tau_{2k}) \in B_2) P^{\Lambda}(\tau_2 \in A_1, \ldots, \tau_{2k} \in A_k).$$

Set $A_1 = A_2 = \cdots = A_k = \mathbb{R}$ to see $P^{\Lambda}(U_s \in B_1,\ U_t \in B_2) = P^0(U(s - \tau_{2k}) \in B_1, U(t - \tau_{2k}) \in B_2)$, thus $\sigma(U_s, U_t)$ is independent of $\sigma(\tau_2, \tau_4, \ldots, \tau_{2k})$.

From part (d) it follows that $\mathbb{E}^{\Lambda} \int_{\tau(2)}^{\tau(4)} |f(U_s) - m|\, \mathrm{d}s < \infty$, so Fubini's theorem yields

$$\mathbb{E}^{\Lambda}\xi_1 = \int_{\mathbb{R}} \mathbb{E}^{\Lambda}((f(U_s) - m) I(\tau_2 \leqslant s \leqslant \tau_4))\, \mathrm{d}s$$
$$= \int_{\mathbb{R}} \mathbb{E}^{\Lambda}(f(U_s) - m) P(\tau_2 \leqslant s \leqslant \tau_4)\, \mathrm{d}s = 0.$$

Since $\{\xi_k,\ k \in \mathbb{N}\}$ forms a sequence of i.i.d. random variables we have for every $k \in \mathbb{N}$

$$\mathbb{E}^{\Lambda}(\xi_1^2) = \frac{1}{k}\mathbb{E}^{\Lambda}\left(\sum_{j=1}^{k} \xi_j\right)^2 = \frac{1}{k}\mathbb{E}^{\Lambda}\left(\int_{\tau(2)}^{\tau(2k+2)} (f(U_s) - m)\, \mathrm{d}s\right)^2$$
$$= \frac{1}{k}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbb{E}^{\Lambda}((f(U_s) - m)(f(U_t) - m) I(\tau_2 \leqslant s, t \leqslant \tau_{2k+2}))\, \mathrm{d}t\, \mathrm{d}s.$$

By Fubini's theorem and by independence of (U_s, U_t) and (τ_2, τ_{2k+2}) this is

$$= \int_{-\infty}^{\infty} \gamma_U(|s - t|) \frac{1}{k} \int_{-\infty}^{\infty} \mathbb{E}^{\Lambda} I(\tau_2 \leqslant s \leqslant \tau_{2k+2}, \tau_2 \leqslant t \leqslant \tau_{2k+2})\, \mathrm{d}t\, \mathrm{d}s,$$

and after the substitution $r = s - t$ for s we get

$$= \int_{-\infty}^{\infty} \gamma_U(|r|) \frac{1}{k} \mathbb{E}^{\Lambda}\left(\int_{-\infty}^{\infty} I(t \in [\tau_2, \tau_{2k+2}] \cap [\tau_2 - r, \tau_{2k+2} - r])\, \mathrm{d}t\right) \mathrm{d}r$$
$$= \int_{-\infty}^{\infty} \gamma_U(|r|) \frac{1}{k} \mathbb{E}^{\Lambda}(\tau_{2k+2} - \tau_2 - |r|)^+\, \mathrm{d}r,$$

where $x^+ := \max\{x, 0\}$. The integrand is bounded by $\mu|\gamma_U(|r|)|$ which is integrable over $\mathbb{R}$ by Lemma 7(c) and by the strong law of large numbers the integrand converges to $\mu\gamma_U(|r|)$ as $k \to \infty$. Thus, by the dominated convergence theorem the claim follows. □

Now we can follow Horváth and Khoshnevisan (1996). Since by Lemma 9(f) $\mathbb{E}^{\Lambda}\xi_1 = 0$ and $\mathbb{E}^{\Lambda}(\xi_1^2) = \mu\sigma^2 < \infty$, the Komlós et al. (1976) approximation proves the existence of a Wiener process $\{\tilde{W}(t),\ t \geqslant 0\}$ on a possibly enlarged probability space such that

$$\sum_{1 \leqslant k \leqslant x} \xi_k - \sigma\sqrt{\mu}\tilde{W}(x) = \mathrm{o}(x^{1/2-\varrho}) \quad \text{a.s. as } x \to \infty. \tag{30}$$

We note that in the original formulation of this construction the Wiener process $\tilde{W}$ approximates the partial sums of some redefined sequence $\{\xi_k'\}$ which has the same distribution as $\{\xi_k\}$. By a coupling argument (cf. Major, 2000, problem 10) we can also find a Wiener process such that (30) holds. For $t > 0$ let $k = k(\omega)$ be such that $\tau_{2k} \leqslant t \leqslant \tau_{2k+2}$. Then we have

$$\begin{aligned}
&\left| \int_0^t (f(U_s) - m)\,\mathrm{d}s - \sigma\sqrt{\mu}\tilde{W}\left(\frac{t}{\mu}\right) \right| \\
&\quad \leqslant \left| \int_0^t (f(U_s) - m)\,\mathrm{d}s - \int_0^{\tau_{2k}} (f(U_s) - m)\,\mathrm{d}s \right| \\
&\qquad + \left| \int_0^{\tau_{2k}} (f(U_s) - m)\,\mathrm{d}s - \sigma\sqrt{\mu}\tilde{W}(k) \right| + \sigma\sqrt{\mu} \left| \tilde{W}(k) - \tilde{W}\left(\frac{t}{\mu}\right) \right| \\
&\quad = I + II + III.
\end{aligned}$$

By Lemma 9(e) and a Borel–Cantelli-type argument

$$\int_{\tau_{2k}}^{\tau_{2k+2}} |f(U_s) - m|\,\mathrm{d}s = \mathrm{o}(k^{1/2-\varrho}) \quad \text{a.s. as } k \to \infty; \tag{31}$$

thus,

$$\begin{aligned}
&\left| \int_0^t (f(U_s) - m)\,\mathrm{d}s - \int_0^{\tau_{2k}} (f(U_s) - m)\,\mathrm{d}s \right| \\
&\quad \leqslant \int_{\tau_{2k}}^{\tau_{2k+2}} |f(U_s) - m|\,\mathrm{d}s = \mathrm{o}(k^{1/2-\varrho}) \quad \text{a.s.}
\end{aligned} \tag{32}$$

as $k \to \infty$. Apply the law of the iterated logarithm to the sequence $\{\tau_{2k+2} - \tau_{2k}\}$ to see

$$|\tau_{2k} - k\mu| = \mathcal{O}(\sqrt{k \log\log k}) \quad \text{a.s. as } k \to \infty, \tag{33}$$

hence $t \sim k\mu$ and with (32) and (30) we get $I = \mathrm{o}(t^{1/2-\varrho})$ and $II = \mathrm{o}(t^{1/2-\varrho})$ as $t \to \infty$. (33) also implies $|t/\mu - k| = \mathcal{O}(\sqrt{k\log\log k}) = \mathcal{O}(\sqrt{t \log\log t})$ a.s., so by the modulus of continuity of a Wiener process (Csörgő and Révész, 1981, Theorem 1.2.1. (1.2.3), p. 30, with $a_T = \sqrt{T \log\log T}$) yields

$$III = \mathcal{O}((t\log\log t)^{1/4}(\log t)^{1/2}) = \mathrm{o}(t^{1/2-\varrho}).$$

It is well known that $W(t) := \sqrt{\mu}\tilde{W}(t/\mu)$ is also a Wiener process and therefore

$$\left| \int_0^t (f(U_s) - m)\,\mathrm{d}s - \sigma W(t) \right| = \mathrm{o}(t^{1/2-\varrho}) \quad \text{a.s. as } t \to \infty.$$

Proof of the Corollaries. If $G = \Lambda$, Corollary 3 follows immediately. If $G = \Phi_\alpha$ for some $\alpha > 0$, i.e., $M_n/a_n \xrightarrow{d} \Phi_\alpha$ then it is well known that $X_k^\# := \log^* X_k$ has distribution function $F^\#(x) = F(e^x)$ for large x and

$$\alpha\left(\max_{1\leqslant k\leqslant n} X_k^\# - \log a_n\right) = \log^*\left(\frac{M_n}{a_n}\right)^\alpha \xrightarrow{d} \Lambda. \tag{34}$$

Since $f(\alpha(\max_{1\leqslant k\leqslant n} X_k^\# - \log a_n)) = g(M_n/a_n)$ for large n we reduced the case $G = \Phi_\alpha$ to the case $G = \Lambda$ and this can similarly be done if $G = \Psi_\alpha$. The norming constants in (34) are not the canonical ones, but under condition (9) the approximation (14) still holds.

For the other corollaries note that for every $f \in \mathscr{M}_0$ there is some $\eta > 0$ with $f \in \mathscr{M}_\eta$. The central limit theorem is now a consequence of (10) and Slutsky's theorem. The laws of the iterated logarithm follow from those for the Wiener process, see Csörgő and Révész (1981, p. 36, Theorem 1.3.1*) for Lévy's LIL and p. 48 Example 2 for Chung's LIL. □

Acknowledgements

I would like to thank a referee for his careful reading.

References

Berkes, I., 1998. Results and problems related to the pointwise central limit theorem. In: Szyszkowicz, B. (Ed.), Asymptotic Methods in Probability and Statistics. Elsevier, Amsterdam, pp. 59–96.

Berkes, I., Horváth, L., 2001. The logarithmic average of sample extremes is asymptotically normal. Stochastic Process. Appl. 97, 77–98.

Blumenthal, R.M., Getoor, R.K., 1968. Markov Processes and Potential Theory. Academic Press, New York.

Breiman, L., 1992. Probability, Classics in Applied Mathematics 7. SIAM, Philadelphia.

Cheng, S., Peng, L., Qi, Y., 1998. Almost sure convergence in extreme value theory. Math. Nachr. 190, 43–50.

Csörgő, M., Révész, P., 1981. Strong Approximations in Probability and Statistics. Academic Press, New York.

Deheuvels, P., 1974. Valeurs extrémales d'échantillons croissants d'une variable alétoire réelle. Ann. Inst. Henri Poincaré Sec. B X, 89–114.

Durrett, R., 1996. Probability: Theory and Examples, 2nd Edition. Duxbury Press, Belmont, California.

Dœblin, W., 1938. Sur deux problèmes de M. Kolmogoroff concernant les chaînes dénombrables. Bull. Soc. Math. France 66, 210–220.

Dwass, M., 1964. Extremal processes. Ann. Math. Statist. 35, 1718–1725.

Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. Modelling Extremal Events. Springer, Berlin.

Fahrner, I., 2000a. An extension of the almost sure max-limit theorem. Statist. Probab. Lett. 49, 93–103.

Fahrner, I., 2000b. Almost sure versions of weak limit theorems. Dissertation Thesis, Universität Ulm; Berichte aus der Mathematik, Shaker Verlag, Aachen.

Fahrner, I., Stadtmüller, U., 1998. On almost sure max-limit theorems. Statist. Probab. Lett. 37, 229–236.

Fahrner, I., Stadtmüller, U., 2000. Strong approximation of maxima by extremal processes. Preprint Universität Ulm.

Feller, W., 1968. An Introduction to Probability Theory and its Applications. Vol. I, 3rd Edition. Wiley, New York, London, Sydney.

Gladysheva, Y., Sakhanenko, A., 1985. Invariance principle for the sum of minima. Siberian Math. J. 26, 677–684 (Translation).

Gosh, M., Jogesh, G., Mukhopadhyay, N., 1975. Almost sure convergence of sums of maxima and minima of positive random variables. Z. Wahrsch. verw. Gebiete 33, 49–54.

Gouet, R., 1989. Embedding in extremal processes and the asymptotic behavior of sums of minima. Statist. Probab. Lett. 8, 219–223.

Grenander, U., 1965. A limit theorem for sums of minima of stochastic variables. Ann. Math. Statist. 36, 1041–1042.

de Haan, L., 1970. On Regular Variation and Its Applications to the Weak Convergence of Sample Extremes. Math. Centre Tracts 32, Amsterdam.

de Haan, L., Hordijk, A., 1972. The rate of growth of sample maxima. Ann. Math. Statist. 43, 1185–1196.

Höglund, T., 1972. Asymptotic normality of sums of minima of random variables. Ann. Math. Statist. 43, 351–353.

Hebda-Grabowska, H., Szynal, D., 1979. An almost sure invariance principle for the partial sums of infima of independent random variables. Ann. Probab. 7, 1036–1045.

Horváth, L., Khoshnevisan, D., 1996. A Strong Approximation for Logarithmic Averages. Studia Sci. Math. Hung. 31, 187–196.

Itô, K., McKean Jr., H.P., 1965. Diffusion processes and their sample paths. Die Grundlagen der mathematischen Wissenschaften, 125. Springer, Berlin.

Komlós, J., Major, P., Tusnády, G., 1976. An approximation of partial sums of independent R.V.'s and the sample DF II. Z. Wahrscheinlichkeitstheorie verw. Gebiete 34, 33–58.

Leadbetter, M.R., Lindgren, G., Rootzen, H., 1983. Extremes and related Properties of Random Sequences and Processes. Springer, Berlin.

Mandl, P., 1968. Analytical treatment of one-dimensional Markov processes. Die Grundlagen der mathematischen Wissenschaften, 151. Academia Prague, Springer, New York.

Major, P., 2000. Coupling methods in probability: the relation between the closeness of random variables and their distribution. Technical Report. Available at http://www.math-inst.hu/ ∼major/probability/couple.html.

Motoo, M., 1959. Proof of the law of the iterated logarithm through diffusion equation. Ann. Inst. Statist. Math. 10, 21–28.

Resnick, S.I., 1987. Extreme Values, Regular Variation, and Point Processes. Springer, New York.

Robbins, H., Siegmund, D., 1971. On the law of the iterated logarithm for maxima and minima. Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability. University of California Press, Berkeley, CA.